

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

On protein - protein search, using sw model

Run on: March 1, 2001, 15:47:02 ; Search time 210.42 Seconds  
(without alignments)  
108.227 Million cell updates/sec

Title: US-09-331-631A-3  
Perfect score: 3532  
Sequence: 1 MAINTSNLCSLLFULLSFL... SSRSTKQQPLVSLIDFVGF 666  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 268485 seqs, 34193795 residues

Total number of hits satisfying chosen parameters: 268485  
Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%, Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_36:\*

1: /SIDS1/gcadata/geneseq/geneseq/AA1980.DAT: \*  
2: /SIDS1/gcadata/geneseq/geneseq/AA1981.DAT: \*  
3: /SIDS1/gcadata/geneseq/geneseq/AA1982.DAT: \*  
4: /SIDS1/gcadata/geneseq/geneseq/AA1983.DAT: \*  
5: /SIDS1/gcadata/geneseq/geneseq/AA1984.DAT: \*  
6: /SIDS1/gcadata/geneseq/geneseq/AA1985.DAT: \*  
7: /SIDS1/gcadata/geneseq/geneseq/AA1986.DAT: \*  
8: /SIDS1/gcadata/geneseq/geneseq/AA1987.DAT: \*  
9: /SIDS1/gcadata/geneseq/geneseq/AA1988.DAT: \*  
10: /SIDS1/gcadata/geneseq/geneseq/AA1989.DAT: \*  
11: /SIDS1/gcadata/geneseq/geneseq/AA1991.DAT: \*  
12: /SIDS1/gcadata/geneseq/geneseq/AA1992.DAT: \*  
13: /SIDS1/gcadata/geneseq/geneseq/AA1993.DAT: \*  
14: /SIDS1/gcadata/geneseq/geneseq/AA1994.DAT: \*  
15: /SIDS1/gcadata/geneseq/geneseq/AA1995.DAT: \*  
16: /SIDS1/gcadata/geneseq/geneseq/AA1996.DAT: \*  
17: /SIDS1/gcadata/geneseq/geneseq/AA1997.DAT: \*  
18: /SIDS1/gcadata/geneseq/geneseq/AA1998.DAT: \*  
19: /SIDS1/gcadata/geneseq/geneseq/AA1999.DAT: \*  
20: /SIDS1/gcadata/geneseq/geneseq/AA2000.DAT: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

8

Result No. Score Query Match Length DB ID Description

Result No.	Score	Query	Match	Length	DB	ID	Description
1	3532	100.0	666	19	W02829		Macadamia integrifolia
2	3412	96.6	666	19	W02828		Macadamia integrifolia
3	3215	91.0	625	19	W02830		Macadamia integrifolia
4	1145	32.4	566	13	R20181		Sequence encoded by Gossypium hirsutum
5	1105.5	31.3	590	19	W02832		Theobroma cacao an
6	1036	29.3	525	19	W02831		Zea mays antimicro
7	895	25.3	593	19	W02835		Hordeum vulgare an
8	875	24.8	637	19	W02837		Peanut allergen Ar
9	873.5	24.7	614	18	W22149		Arachis hypogaea a
10	873.5	24.7	614	19	W02834		Peanut allergen A
11	865.5	24.5	626	20	Y15244		Peanut allergen 11
12	865.5	24.5	626	20	Y25657		

9

ALIGMENTS

RESULT	1
ID	W02829 standard; Protein; 666 AA.
XX	
AC	W02829;
XX	
DT	27-OCT-1998 (first entry)
XX	
DE	Macadamia integrifolia antimicrobial protein.
XX	
KW	antimicrobial protein; infestation; control.
XX	
OS	Macadamia integrifolia.
XX	
FH	Location/qualifiers
FT	peptide 1..28
FT	protein /note= "signal peptide"
FT	protein 29..666 /note= "mature protein"
XX	
PN	W09827805-A1.
XX	
PD	02-JUL-1998.
XX	
PF	22-DEC-1997; 97WO-AU00874.
XX	
PR	20-DEC-1996; 96AU-0004225.
XX	
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX	
PI	Bower NI, Goulet KC, Green JL, Manners JM, Marcus JP;
XX	
DR	WPI: 1998-377279/32.
XX	
N-PSDB:	V42311.
XX	

Peanut allergen Ar  
Glycine max anti.  
Soybean beta-congl  
G. max SBP protei  
G. max SBP2 protei  
G. max truncated S  
Ara h 1 allergen p  
G. max truncated S  
Phaseolin A chain  
T. gondii immuno  
A human trichohyal  
HHV8 ORF 73 Protei  
Leucocytozoan prot  
Human thyrotropin  
P. falciparum LSA 9  
Human follicle sti  
Modified oat globi  
Rice storage prote  
Peanut allergen, A  
Ara h 3 allergen s  
P. falciparum LSA 9  
Human STE20-relate  
Amino acid sequenc  
Human CPC28-#3 RNA  
Protein regulating  
Full length human  
Human ZC3 protein.  
Human GEK2\* protein.  
Biorhythm marker p  
Human ZC2 protein.  
LexA/NUMA fusion p

PT	Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals
XX	antimicrobial protein; infestation; control.
PS	Claim 1; Page 39-41; 96pp; English.
XX	The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian animals.
CC	Sequence 666 AA:
XX	
QQ	Query Match 100 %; Score 3532; DB 19; Length 666; Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MAINTNSLCSLFLISLFLISTVSLAESERDQEYECKRQCMLETSQMRRCSQCD 60
Db	1 maintnslcslflislfislistvslaeseidrqyeckrqcmlqletsqmrrcsqd 60
QY	61 KRFEEDIDWSKYDNQDPDQTCQCQCORRCQESPRQOYCQRCKEICEEEENRQR 120
Db	61 krfeedidwskydnqdqptocqqcqrcciqesprqyqycqrkrcikeeeeeeynrq 120
QY	121 DPPQQEEOCQECORETEREPHRMOTQCORRCERRYEDEKRKQKREEQQDEEYERM 180
Db	121 dppqqeqecoreterephrmocqcorrcerryekekraqkryeqqpredeekyerm 180
QY	181 KEEDNKRDPDQPREIDCCRRCBQERQHQOCORRCQRCRQDORHGRGDLINPQRGSGRK 240
Db	181 keednkrdpdqpreyedorrccrqeqprqyqccrqcraqgqhqggdlnpqrgsgry 240
QY	241 EEEGEKQSNDNPYFDERSLSIRFRTEEGHSVYLENFGRSKULRNKRYLVLEANPNA 300
Db	241 eegeekqsndnpypfderlsstrfrteeghsvlenfygrsklrlaknkylvleanpna 300
QY	301 FVLPTILDAATLLVGGRAKLMHRDRNRESYNLCGDVIRIPAGTTFLINRNNERL 360
Db	301 fvlpthldadaillyvggralkmihndresynlcgdviripagttflinrnnrl 360
QY	361 HIAKFLQTISPGQKKEFPAGQGNPEPISTFSKTEALEALNTQAEERLGVLGOREGV 420
Db	361 hiakflqtispqgkkefpagqgnpeplstfsktelealntqaerlgvlgoaregv 420
QY	421 ITSASEQIRELTDDSESRWHRTRGGESSRGPYNUFNKRPLYSNKYGADEVKPEDYR 480
Db	421 iisaseqireltddsesrtwhirrgessrgpynlknkrlplysnkygqayevkpedyr 480
QY	481 OLQDMWSVFIANITQSMGPFFNRSTKWWVASGEADYEMAPHLSRHGGGGKR 540
Db	481 q1qlqdmwsvfianitqsmgpffnrstkkwwasgeadymaphlsrhggggkr 540
QY	541 HEEEVDHYEQVKARLSKREATWVPGHPPVVVFSSGENNLIFARGINAQNHNENFLAGR 600
Db	541 heeevdhyeqvkarskreatwvpghpvvfssgennlifarginaqnhenflagr 600
QY	601 ERNVLOQIEPQAMELAFAAPRKEVEELFNQODESTFFPGKPHQHQSRSRSTKQQPLVST 660
Db	601 ernvloqiepqamealaaprkveelfnqsdesifppqphqgssrstkqqplvsi 660
QY	661 LDFVGF 666
Db	661 ldfvgf 666
RESULT	2
W62828	W62828 standard; Protein; 666 AA.
XX	W62828;
XX	27-OCT-1998 (first entry)
DE	macadamia integrifolia antimicrobial protein
KW	
OS	Macadamia integrifolia
XX	
KEY	Location/Qualifiers 1..28 /note= "signal peptide" 29..666 "mature protein"
FT	peptide
FT	protein
FT	
XX	
PN	W09827805-A1.
XX	
PD	02-JUL-1998.
XX	
PF	22-DEC-1997; 97WO-AU000874.
XX	
PR	20-DEC-1995; 96AU-0004275.
XX	
PA	(RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.
XX	
PI	Bower NI, Goulier KC, Green JL, Manners JM, Marcus JP;
XX	
DR	DR N-PSDB; V42310.
XX	
PT	Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals
XX	
PS	Claim 1; Page 34-36; 96pp; English.
XX	
CC	The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian animals.
CC	
XX	
SQ	Sequence 666 AA:
Query Match 96 %; Score 3412; DB 19; Length 666; Best Local Similarity 96.1%; Pred. No. 1.5e-298; Mismatches 14; Indels 0; Gaps 0;	
QY	1 MAINTNSLCSLFLISLFLISTVSLAESERDQEYECKRQCMLETSQMRRCSQCD 60
Db	1 maintnslcslflislfislistvslaeseidrqyeckrqcmlqletsqmrrcsqd 60
QY	61 KRFEEDIDWSKYDNQDPDQTCQCQCORRCQESPRQOYCQRCKEICEEEENRQR 120
Db	61 krfeedidwskydnqdqptocqqcqrcciqesprqyqycqrkrcikeeeeeeynrq 120
QY	121 DPPQQEEOCQECORETEREPHRMOTQCORRCERRYEDEKRKQKREEQQDEEYERM 180
Db	121 dppqqeqecoreterephrmocqcorrcerryekekraqkryeqqpredeekyerm 180
QY	181 KEEDNKRDPDQPREIDCCRRCBQERQHQOCORRCQRCRQDORHGRGDLINPQRGSGRK 240
Db	181 keednkrdpdqpreyedorrccrqeqprqyqccrqcraqgqhqggdlnpqrgsgry 240
QY	241 EEEGEKQSNDNPYFDERSLSIRFRTEEGHSVYLENFGRSKULRNKRYLVLEANPNA 300
Db	241 eegeekqsndnpypfderlsstrfrteeghsvlenfygrsklrlaknkylvleanpna 300
QY	301 FVLPTILDAATLLVGGRAKLMHRDRNRESYNLCGDVIRIPAGTTFLINRNNERL 360
Db	301 fvlpthldadaillyvggralkmihndresynlcgdviripagttflinrnnrl 360
QY	421 ITSASEQIRELTDDSESRWHRTRGGESSRGPYNUFNKRPLYSNKYGADEVKPEDYR 480



PT expression vectors

XX  
PS Claim 4; Fig 2; 59pp; English.

CC The inventors claim a 67 kD and 31 kD T. cacao protein, and fragments, and encoding DNAs. The 47 kD and 31 kD proteins are derived from the 67 kD precursor. T. cacao protein cDNA was detected in a cDNA library prepared from immature cocoa beans RNA using a probe based on the AA sequence of a CNBr peptide common to the 47 kD and 31 kD polypeptides. Homology searches revealed close homologies between the 67 kD polypeptide and the vicilins, which are seed storage proteins.

CC  
XX  
SQ Sequence 566 AA;

Query Match 32.4%; Score 1145; DB 13; Length 566;  
Best Local Similarity 41.1%; Pred. No. 1.4e-94;  
Matches 237; Conservative 107; Mismatches 175; Indels 58; Gaps 13;

QY 109 ICEEEEFYINR--QRDQQQVEQCQCRCORHETEPHMQRQORERRYEKEKRQQRY 165  
:|| :|| | :||:||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 22 lcsqvsaygrkqyeprpqyegqrceateareeqeqracer-----ey 70

QY 166 EBDQRIIDEDEKVERMKEDDNKRDPDQEYEDCRRCEEQE--PRQYOCORCRRCRQHQH 223  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 71 keqqrqgeel-----qrqyqgcqgreqgqgreqcqkrcwqeqyek 116

QY 224 GRGGDLINPQRGSGRYEEGEEKQSKPNPYFDE-NSLSTFRTEGHISVLENFGRSL 282  
| :|| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 117 er-gehenehyhhkkkrseeeeqggrrnppypkrstfqtrrdeeqnfklqrfaenssp 175

QY 283 LRALKWYRLVLEANNAFLNPAVLPQLDADAILVTGGRGALKMWHDRNRESYNLEGDVIR 342  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 176 lkgindylarlafeanpntfiphqdaaifyvtngktitfthenkessynvqgtvus 235

QY 343 IPAGTTPYFLIMRNDRNLRHAKFLQTISTPQYKEFPAGCONPEYLSTPSKELIAAL 402  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 236 vpagstvvysqsdngqekltiavlapvnspksgyefpagnkpsyygafsvyletvf 295

QY 403 NTQAERLURGVIGQQR-----EGYITISASOEORELTRDSESRWRHIRGGSSRG 454  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 296 ntqrekleileeprqkrqgqkqmrakpeqkiraqsdgatbspr---hrgg---rla 349

QY 455 YNLFRKRPLSNKYKSOAYEVKPEDRQLQMDVSFIANTIQGSMAGPFNTRSKVWW 514  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 350 inllsqspvysnqngtffacepdtsqfqmdavasfakingqaiifvphynskatfvu 409

QY 515 ASGEHALVEMAPPHLSRKHGGRGGK--RHEBEDD-----HYEWKARLSKRE 566  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 410 tdygyraqmacphlsrsqssqsgqdrqeqeeseeetgefqqykaplspgdfrvpa 469

QY 567 GHPVWVVSSENLLIFAFGNGNAONHNHENFLAGRNRVLOQIEPAMELAFAAPPKEVEE 626  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 470 ghavttfaskoqgpihnavaflqnaqnqriflagk.knlvrqmdaseakelsrgvpkskv 528

QY 627 LFNSQDESIFPPGPQHQQSSRSRKQQPLVISDF 663  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 529 ifnpndesyfmsfsqqr---rdergnplasidt 562

OS Gossypium hirsutum.

XX  
PN WO9827805-A1.

XX  
PD 02-JUL-1998.

XX  
PF 22-DEC-1997; 97WO-AU00874.

XX  
PR 20-DEC-1996; 96AU-0004275.

XX  
PA (RETR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.

PI Bower NI, Gouitter KC, Green JL, Manners JM, Marcus JP;

DR DR; 1998-377279/32.

PT Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals

PT  
XX  
PS Claim 1; Page 49-51; 96pp; English.

CC The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian animals.

XX  
SQ Sequence 590 AA;

Query Match 31.3%; Score 1105 5; DB 19; Length 590;  
Best Local Similarity 38.4%; Pred. No. 5.5e-91;  
Matches 233; Conservative 113; Mismatches 174; Indels 87; Gaps 14;

QY 76 DDPQTICQCQCORCRQCROQESGRQQVCQRCKEICEEEEFINRQR--DPOQYEQCQRC 133  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 35 ddpkkyedcrccrdtrgqkqeqgqceccsksqygekdqqrhpdpqarryecqc 94

QY 134 ORHETPRHOMCOOCERREKERKQQRYEEQOREDEEKEYEERMKEDDNKRDPDQE 193  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 95 r-qqeerqqqeqcqrckifkeeqeqq-----sqrq 123

QY 194 YEDCRRCEDQE--PRQYOCQRREQQRQGRGGDLINPQRC-----GSGVEEGEE 245  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 124 fqeqggchqchgeqrpkqeqvcrecrekyqe-----npwgrrereeeeeteeqq 175

QY 246 KQSDNPYYFDBRSLSTRFRBEGHISVLENFYGRSKLRLAKNVRVLLEANPNAFLVPT 305  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 176 eqshpofhfhnsrfsqsrfreeqhnfnvlqtfaspriplqrgnefrsileanpntfvlph 235

QY 306 RUDADAILVPGGRGALKMWHDRNBSYNECQDGIVRIPAGTFILNRONERIHKAF 365  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 236 hodaekiyitngtngtlfthenkessynvpgvvvkvpagstylandnkekliavl 295

QY 366 LOTISPQGYKEFPGAGQPEPYLSTESTESEILEALNTOAERLURGVUG-----QORE 418  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 296 hrpvnnpgqfeeffpagsqrpqsyiratsreilepafntrseqidelfggrqsrqgqg 355

QY 419 GVIIISQEQRELTRDSSRRWIRRGGESSRCPYNIENRKPJISNKYGADEVKPED 478  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 356 qmfrkasqedqraltsqeqatpr---ekgee--tafnllsqtprysngqrfteacpe 409

QY 479 YROLQDMDSFIANTIQGSMAGPFNTRSKVWWVASQHADVENACPHLSGRHGRGC 538  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 410 frqlrdinvtsalqnlqgsifvphynskatfvltvegnyyaenvsphlp---rqss 464

QY 539 KRHEEEDV-----HYEWKARLSKRE 585  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

QY 465 yeeeeedeedeqedeerqgqykirslsrgdfivpantpvtfvaqsqnqlmtgf 524

QY 586 G-----INQNHENFLAGRNRVQIIEQAMELAFAAPRKEVELFSQDESIFFPGP 640  
:||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||  
Db 525 glynqninphnqgrifvagknvh-rqwdqskakelafgvssrvidefnspqesyfvs- 582

QY 641 RQHQOQS 647

RESULT 5  
W62832  
ID W62832 standard; Protein: 590 AA.  
AC W62832;  
XX DT 27-OCT-1998 (first entry)  
XX DE Gossypium hirsutum antimicrobial protein.  
XX KW antimicrobial protein; infestation; control.  
XX



Db 260 rlgfrgngdqkqgivrateeqtrellrrhastegghphwplppge srgyslildgrpsia 318  
 Qy 466 NKGQAYEVKPDYRQLQDMVSFTANTIQGSMGPPFNRTSKVWVAGADEVEMAC 525  
 Db 319 nqhgqlyeadrsfhdlaehovsvsanitgsmasplntsrsklayvpqkgraeiv 378  
 Qy 526 PHLSGRHG---GRRGCKRHEEEEDH-----YEQVKARLSKREATWVPHCPV 571  
 Db 379 phrqsgqgesererdkgrrseeeeeseeqeaqqgyhtiarispgtafvvpaghpfv 438  
 Qy 572 FVSSGENNLIFARGINAQNHHENFLAGRERNVLQIEPQAMELAFAPRKEVELENSQ 631  
 Db 439 avasdsnqlqvcfahdrnkvflagdnvldqkldrvalstaskaeedevlgsr 497  
 Qy 632 DESIFPGPRO--HQOQSRSRSTKQD 655  
 Db 498 rekgtfpqpeesggheereqeeere 524

RESULT 8  
 W62837 standard; Protein; 637 AA.  
 XX W62837;  
 AC W62837;  
 DT 27-OCT-1998 (first entry)  
 DE Hordeum vulgare antimicrobial protein.  
 KW antimicrobial protein; infestation; control.  
 OS Hordeum vulgare.  
 PN W09827805-A1.  
 XX PD 02-JUL-1998.  
 PF 22-DEC-1997; 97WO-AU00874.  
 PR 20-DEC-1996; 96AU-0004275.  
 XX (RETR) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 XX Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;  
 DR WPI: 1998-377279/32.  
 PT Novel anti-microbial protein from e.g. Macadamia integrifolia - useful for controlling microbial infestations of plants or mammals  
 PS Claim 1; Page 60-62; 96pp; English.  
 XX The sequence is that of an antimicrobial protein which can be used to control microbial infestations in plants and mammalian animals.  
 CC Sequence 637 AA;

Query Match 24.8%; Score 875; DB 19; Length 637;  
 Best Local Similarity 35.6%; Pred. No. 3.5e-10;  
 Matches 227; Conservative 99; Mismatches 206; Indels 106; Gaps 22;

Qy 70 SKYDDNQDPQ-TDCQOCQRCCRQDQESGPROQQYCORCCEICEPEEEYNQRDQOQYE 127  
 Db 27 ashdeddrqgqhsaqvgcrqrer-pr-ysharcqec-----rdqgqh- 71  
 Qy 128 QCQERCORHETEPRIMOTCQRCRERYEKRKQKRYEQOREDE---KYSERMKE 182  
 Db 72 -----grhqeegqgrgwghegedereehgqgrhgegeeehgrgrgrhgege 125  
 Qy 183 EDNKADPOQRREYEDRRRCQEQPQOYCQRCRQDQHGRG---GDLINPORGGSG 238

Db 126 eergry-----hgrhage-----reererggrgrgegeeeeegrrg 165  
 Qy 239 RVEEGE--EKOSDN--PYFVERSLSPFRFREGHSVYLEFYGRSKLRLAKNLYLVL 294  
 Db 166 rigegeerdeegdstrrryvfpstrriqdgforalrfdqgvslrliridrvain 225  
 Qy 295 EANPNAFVLPLPHLDADDAILLYTGGRGALKMTHRDNRSYNEECGDVIRAPGTTFLINR 354  
 Db 226 evprafvwpvptdadggyvqaqgegyvltvengekrsytkegdyivapagsimlant 285  
 Qy 355 DNNERLHTIAKLIQTI3TPGQKEFPAGGQNEPYLSTSFEILEAALNTAQERLRCVLG 414  
 Db 286 dgrrkliakihtisvpgkf-qfl---svkpklasiskrvlaakfksderlerfn 339  
 Qy 415 QO-----REGVIISASQEORLTDSE--SRWHHIRGGESSRGPNLFNKRPLYS 465  
 Db 340 arqgqektrssivraseeqeirlrealeaaegqghnwplphrgsrdtfnleirpkia 399  
 Qy 466 NKGQAYEVKPDYRQLQDMVSFTANTIQGSMGPPFNRTSKVWVAGADEVEMAC 525  
 Db 400 nrhgrlyeadarsfhalanqvravvanitqsgmstapyintqsfklawvliegegeqvic 459  
 Qy 526 PHL----SGRHGRRGKRMHEEE-----VHEYEQVKARLSKRE 560  
 Db 460 phlgreserehbg---grrreecedqrgqrrgyseseseeeqryetvavrsvrs 517  
 Qy 561 ATVPVGHGVVFVSS--GNENLIFARGINAQNHHENFLAGRERNVYLQQPAMELAF 618  
 Db 518 afvvppghpvveissssggssnlqvvcfeineernernevwlacr-nvvigkqspagltfg 576  
 Qy 619 APRKEVEELFNSQDES-FFGPQPROHQOQSRSRSTKQD 655  
 Db 577 rparevdevfaqdgdegfvaaq----eqqseqeqeq 610

RESULT 9  
 W22149 standard; Protein; 614 AA.  
 XX W22149;  
 AC W22149;  
 XX DT 29-DEC-1997 (first entry)  
 DE Peanut allergen Ara h1.  
 XX KW Peanut; seed storage protein; allergen; allergy; hypersensitivity; monoclonal antibody; ELISA; analysis; Ara h1.  
 OS Arachis hypogaea strain Florunner.  
 XX FH Key Location/Qualifiers  
 FT Peptide 1..22  
 FT Protein /label= Sig\_peptide  
 FT Modified-site /label= Mat\_protein  
 FT /note= "N-glycosylation site"  
 XX FN W09724139-A1.  
 XX PD 10-JUL-1997.  
 XX PF 23-SEP-1996; 96WO-US15222.  
 XX PR 04-MAR-1996; 96US-0610424.  
 XX PR 29-DEC-1995; 95US-0009455.  
 XX (UYAR-) UNIV ARKANSAS.  
 RI Bannon GA, Burks AW, Cockrell G, Helm RM, Stanley JS;







PT Novel anti-microbial protein from e.g. Macadamia integrifolia -  
 XX useful for controlling microbial infestations of plants or mammals.  
 PS Claim 1; Page 63-65; 96pp; English.

XX The sequence is that of an antimicrobial protein which can  
 CC be used to control microbial infestations in plants and mammalian  
 CC animals.

XX Sequence 605 AA;

Query Match 24.2%; Score 855.5; DB 19; Length 605;  
 Best Local Similarity 31.4%; Pred. No. 1.8e-68;  
 Matches 194; Conservative 143; Mismatches 211; Indels 69; Gaps 15;

QY 72 YDNQDPDQTCQQCQRCRQESGRQQYCQRC-----KECEEE--EVNRQDP 122  
 Db 27 ywkeenkphn--kclqscnsrdsyrsnq-a-charcnlkveekeecegeiprpprppqip 83

QY 123 QQYEQCQERQRHETEPRIQTCQRCRREKEKRKQOQRYYEROOREDEEKYERMK 182  
 Db 84 erepqapgekeeededeqrpripfpipfqprqprqeeehqeqewpr-keekgkseeded 143

QY 183 EDDNKRPOORREYEDCRRRSCQOOPRQYOCORRCRQEQORHGGREGDLINPORGSSGRYE 242  
 Db 144 edeeqderqfpip--rphphkeerneeedd-eeqres-----aa 181

QY 243 GEEKQ----SDNPYFDESSLSTRFRTEGHISYLENFGRSKLRAKNYLVEAN 297  
 Db 182 sdeselrrrhkknknfifgsnorfeflfknqygrirvifqfrqrsqqlqn1rdyrilefnk 241

QY 298 PNAFWIPLHDADATLVTGGRGALKMHRDNRESYNLEQDVIRIPAGTTFLINRDNN 357  
 Db 242 pntlliphadadyivilingtallsvnddrdsyrlsqdalrvpsgttlyyvnpdn 301

QY 358 ERHLIAKFLOTISTPQYKSKFPAGCONPYPRLSTFSKELEALANTQEBRLGVV--- 413  
 Db 302 enrilitlaipvnkpgrfesfflsssteagqsyigfsrnileasydtkfeenkvlfsre 361

QY 414 ----GOQR--EGVITASASBQIRELTDRDSESRRWHRIGGESESRGPYNLFNRPYLN 465  
 Db 362 egqqgagqrqiesviveiskeqralskrakssrti...-sseakptnrlrsropysn 417

QY 467 KYGQAYEVKEDPYQLQDMMSVNFANITOQSMMPFFNTRSTKVVVVASGEADVEMACP 526  
 Db 418 klgkfffeitpeknpglrlidifilisvdmmegaldiphfnskavilvinegdanielv-- 475

QY 527 HLSGRHGG--RGGRRHEEEDVHYEQVKRSLRSKEAIVVPGHVVVFVSSGENLLFA 584  
 Db 476 -----glkqeqqeqqeqqplervkyraelseqadifivipagqypvv-natsnlifa 526

QY 585 FGINAQNHENFLAGRERNYILOQEPOMALAFAAAPRKVEELNSQDESIFFPGPROQ 644  
 Db 527 iginaenqnrflagsqndvisqipsqvgelafpsaqavekliknqresyfvdaqpk 586

QY 645 QQSRSRSTKQQQPLVSL 661  
 Db 587 eeqnkgrk--qpissil 601

RESULT 14  
 W62838 standard; Protein; 605 AA.  
 XX  
 XX W62838;  
 XX  
 DT 27-OCT-1998 (first entry)  
 DE Glycine max antimicrobial protein.  
 XX  
 KW antimicrobial protein; infestation; control.  
 XX  
 OS Glycine max.  
 XX  
 PN W09827805-A1.

PD 02-JUL-1998.  
 XX  
 PF 22-DEC-1997; 97WO-AU00874.  
 XX  
 PR 20-DEC-1996; 96AU-0004275.  
 XX  
 PA (REFR-) COOP RES CENT TROPICAL PLANT PATHOLOGY.  
 XX  
 PI Bower NI, Goulter KC, Green JL, Manners JM, Marcus JP;  
 XX  
 DR WPI: 1998-377279/32..  
 XX

RESULT 15  
 Y40999 standard; protein; 605 AA.  
 XX  
 XX  
 AC Y40999;  
 XX  
 DT 06-DEC-1999 (first entry)  
 XX  
 DE Soybean beta-conglycinin protein sequence.  
 XX  
 KW Peanut; allergen; Ara H 1; IgE; immunoglobulin E; epitope; Ara h 3;  
 KW allergic reaction; soybean; beta-conglycinin.

OS Glycine max.  
 XX  
 PN WO9945961-A1.  
 XX  
 PD 16-SEP-1999.  
 XX  
 PF 12-MAR-1999; 99WO-US05494.  
 XX  
 PR 12-MAR-1998; 98US -0077763.  
 PR 11-MAR-1999; 99US -0077763.  
 XX  
 PA (YUAR-) UNTV ARKANSAS.  
 XX  
 PI Burks W, Helm RM, Cockrell G, Bannon GA, Stanley JS, Shin DS;  
 PI Sampson H, Compadre CM, Huang SK, Maleki SJ, Kopper RA;  
 XX  
 DB warr. 1999-551129-146

QY 523 MACPHLISGRHGG--RKGKRHEEEEDVHYEOKVARKLSKREAVVPGHPPVVFVSGNENL 580  
 Db 474 lv-----gikqeqgqeqgeeeplvrvkyraelseqdifvibagypvvv-natsnl 522  
 QY 581 LLFARGINAONNHENFLAGRNVLQIERQAMELAFAAFRAKEVEELPNSODESFPPGP 640  
 Db 523 nffaiginaenqrflagsqdnvtsqipsqsvqelafpgsaqavkkilkqrksyfvdq 582  
 QY 641 RQHQOQSRSRKQQQPLVSL 661  
 Db 583 pkkkegnkpk - gpissl 601

Search completed: March 1, 2001, 15:47:07  
 Job time: 232 sec

